Pure

Syddansk Universitet

Spectrogram analysis of mechanical events in sound production of animals

Elemans, Coen; Muller, Mees; Heeck, Kier

Published in:
Bioacoustics

Publication date:
2008

Document Version
Submitted manuscript

Link to publication

Citation for pulished version (APA):
Elemans, C., Muller, M., & Heeck, K. (2008). Spectrogram analysis of mechanical events in sound production of animals. Bioacoustics, 18, 183-212.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 16. jan.. 2017

Bioacoustics
The International Journal of Animal Sound and its Recording, 2008, Vol. 18, pp. 183–212
© 2008 AB Academic Publishers

SPECTROGRAM ANALYSIS OF ANIMAL SOUND PRODUCTION

COEN P.H. ELEMANS[1*], KIER HEECK[2] AND MEES MULLER[1]

[1]*Experimental Zoology Group, Wageningen University, Marijkeweg 40, 6709 PG Wageningen, The Netherlands*
[2]*Condensed Matter Physics, Division of Physics and Astronomy, Free University of Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands*

ABSTRACT

Spectrograms visualise the time-frequency content of a signal. They are commonly used to analyse animal vocalisations. Here, we analyse how far we can deduce the mechanical origin of sound generation and modulation from the spectrogram. We investigate the relationship between simple mathematical events such as transients, harmonics, amplitude- and frequency modulation and the resulting structures in spectrograms. This approach yields not only convenient statistical description, but also aids in formulating hypotheses about the underlying mathematical mechanisms. We then discuss to what extent it is possible to invert our analysis and relate structures in spectrograms back to the underlying mathematical and mechanical events using two exemplary approaches: (a) we analyse the spectrogram of a vocalisation of the Bearded Vulture and postulate hypotheses on the mathematical origin of the signal. Furthermore, we synthesise the signal using the simple mathematical principles presented earlier; (b) we use a simple mechanical model to generate sounds and relate experimentally observed mechanical events to characteristics of the spectrogram. We conclude that although knowledge of sound producing systems increases the explanatory power of a spectrogram, a spectrogram *per se* cannot present unambiguous evidence about the underlying mechanical origin of the sound signal.

Keywords: Bioacoustics, biomechanics, Fourier analysis, *Gypaetus barbatus*, Duffing equation

INTRODUCTION

Throughout the animal kingdom a wide range of biological mechanisms of generating sound have evolved. Sound production by animals is often based on complex mechanical events, ranging from cavitation induced sound by shrimps (Versluis *et al*. 2000) to the energy conversion of flow to vibrating structures in the larynx of mammals (e.g. Paulsen 1967). The generated sound amplitude can show up as

*Corresponding author. Current address: Institute of Biology, University of Southern Denmark, Campusvej 55, DK-5320, Odense M, Denmark. Email: coen@biology.sdu.dk

a clear periodic function of time. In many cases, however, a sound possesses a seemingly aperiodic character (e.g. a sound recording of a piece played by a symphony orchestra; see e.g. Von Békésy (1960)). Although in the time domain hardly any structure can be discovered, the frequency domain may reveal a clear composition of periodic signals. Spectral analyses to study the frequency domain are widespread in science and engineering with applications ranging from econometrics and astronomy to Geographical Information Systems (GIS), Remote Sensing (e.g. Lilisand & Kiefer 2000) and acoustics.

A convenient way to visualise a sound recording in the frequency domain is a spectrogram, which shows the spectral composition (e.g. amplitude) of a time window sliding over the signal. Spectrograms based on the Fourier Transform are widespread in biological sciences such as animal behaviour and bioacoustics. Their value (Hall-Craggs 1979), trade-offs and limitations (e.g. Beecher 1988; Bradbury & Vehrencamp 1998; Clark *et al.* 1987; Cohen 1995; Spiegel 1974; Williams & Slater 1991) have been well studied. Other spectral analysis techniques, such as zero-crossing analysis (Staddon *et al.* 1978), wavelets (Wakeling *et al.* 2002), spectral derivatives (Tchernichovski & Mitra 2000), optimal kernel designs (Jones & Baraniuk 1995), or others (Darden *et al.* 2003; Gardner & Magnasco 2005, 2006; Mbu Nyamsi *et al.* 1994) are used to construct different time-frequency representation with e.g. a higher time-, and frequency resolution. Various techniques have been developed to correlate sound structure in a spectrogram (e.g. Clark *et al.* 1987; Cortopassi & Bradbury 2000; Khanna *et al.* 1998) with various contextual variables, such as habitat, social context, and inter- and intraspecific differences.

The Fourier transform decomposes any infinite signal in sine waves with certain amplitude (or power), frequency and relative phase. In a normal spectrogram, only the amplitude (or power) of the FFT analysis is displayed, also called the *amplitude* spectrogram (Léonard 2000). The relative phase information of the FFT can be equivalently used to construct a phase-time representation, which is called a *phase* spectrogram (Léonard 1997, 2000; Léonard *et al.* 2000). To our knowledge, the phase spectrogram has not been used in published animal sound studies.

The spectrogram is used to present the time-frequency information of a signal, but what additional information can we extract from the spectrogram? If we are interested in the physiological production process of the sound signal, what information does a spectrogram contain on the mechanical origin of the signal? In the present paper, we investigate the relationship between simple mathematical events and the resulting structures in spectrograms. Furthermore, we explore whether we can invert our analysis and relate structures in spectrograms back to the underlying mathematical and mechanical events. First, we will start by studying the shapes and structures

in spectrograms resulting from a selection of mathematically defined sound signals. We then discuss extent to which it is possible to invert our analysis and relate structures in spectrograms back to the underlying mathematical and mechanical events using two exemplary approaches: (a) we analyse the spectrogram of a vocalisation of the Bearded Vulture *Gypaetus barbatus* L. and postulate hypotheses on the mathematical origin of parts of the signal. We subsequently synthesise the signal using the simple modulation and manipulation principles presented earlier; (b) we use a simple mechanical model to generate sounds and relate experimentally observed mechanical events to characteristics of the spectrogram. The latter two approaches are used sparsely in bioacoustics, but have proven to be of great help in directing working hypotheses in other fields of biology. The last section contains our conclusions.

From signal generation to spectrogram. The appearance of simple mathematical signals in spectrograms

Extensive terminology has been developed to describe the most common sound features in spectrograms of animal vocalisations. Examples found in literature are 'pure tone', 'harmonic stacks', 'broadband pulse', 'amplitude- and frequency modulation', 'sidebands' and 'noise'. To understand what kind of signals can cause these structures, we will start by studying the shapes and structures in spectrograms that are the result of a selection of mathematically defined sound signals. This approach provides a basis to interpret signals of which the exact composition and mechanical origin is not *a priori* known. We will generate some of these common features by simple permutations of sine waves in the time domain. The advantage of mathematically defined signals is that the emergence of spectral features can be explained without ambiguity due to error or variation introduced by measurements. To increase readability, we have included the mathematical derivations in the Appendices and not in the main text. We will not provide a detailed explanation on how to construct a spectrogram, because there are many excellent descriptions available in the literature (e.g. Bradbury & Vehrencamp 1989 and many tutorials). In this paper, to construct spectrograms we will use digital Fourier transforms (dFFT), the most widespread and easy to use transform used in bioacoustics.

Harmonics

Harmonics are integer multiples of a base frequency or fundamental frequency (see Figure 1). All non-sinusoidal, but periodic, time-

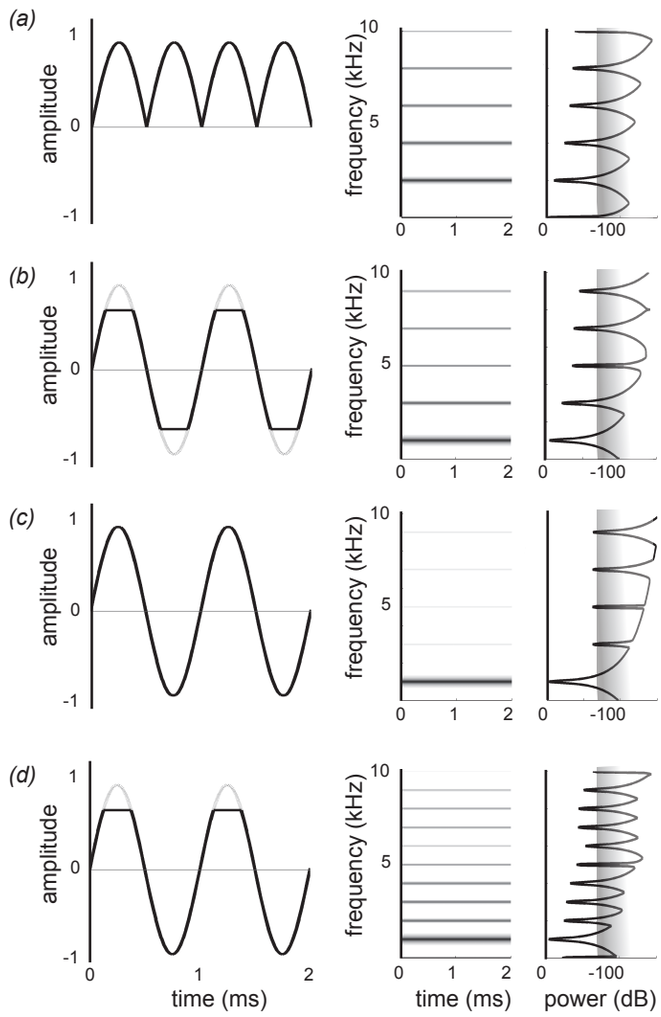

Figure 1. Effects of nonlinearity on sinusoidal time-signals. From left to right: the time signal and associated spectrogram and amplitude spectrum. The shaded bars in the amplitude spectrum indicates the edge of the displayed dynamic range in the spectrogram. *(a)* Full wave rectified sine wave. The fundamental frequency doubles and both odd and even harmonics occur. *(b)* Sine wave clipped symmetrically at A_c = 0.7·A. Only odd harmonics occur in the power spectrum. *(c)* Sine wave clipped symmetrically at A_c = 0.99·A. Even in the latter case, harmonics occur. *(d)* Sine wave clipped asymmetrically at A_c = 0.7·A. Both odd and even harmonics occur. ω_c = 1 kHz; F_s= 40 kHz, *nfft* = 1024.

functions generate harmonics in their spectrum (e.g. Bradbury & Vehrencamp 1998). Because perfect sine waves are rare in real world signals, harmonics are ubiquitous. In some special cases, it is possible to calculate analytically the harmonic content of a signal in a Fourier series (see Appendix B1). The resulting parameter values provide a convenient statistic for comparison and are used in e.g. biomechanics locomotion literature (e.g. Askew & Ellerby 2007).

To generate harmonics as a mathematically defined signal, we use the following example. We consider the absolute value of a sinusoidal signal, also referred to as 'full wave rectification' of a simple sine wave:

$$y(t) = \left|A \cdot \sin\left(\omega_s \cdot t\right)\right| = A \cdot \left[\frac{2}{\pi} - \frac{4}{\pi}\left(\frac{\cos\left(2 \cdot \omega_s \cdot t\right)}{1 \cdot 3} + \frac{\cos\left(4 \cdot \omega_s \cdot t\right)}{3 \cdot 5} + \frac{\cos\left(6 \cdot \omega_s \cdot t\right)}{5 \cdot 7} + \ldots\right)\right] \quad \text{Eqn. 1}$$

Here, ω_s is the angular frequency of the signal, and A is the amplitude. The frequency (in Hz) of the signal is $f = \omega_s/2\pi$. The time signal (Figure 1a) can be expanded into a series consisting of multiples of $2\omega_s$: [cos $(2\omega_s\ldots)$ + cos $(4\omega_s\ldots)$ + … in eqn (1)]. The same expansion of a cosine function also shows only multiples of $2\omega_s$ (see Appendix B1). Full wave rectifying thus results in a fundamental frequency of $2\omega_s$ and not ω_s. The spectrogram (centre column in Figure 1) contains a harmonic stack with odd and even harmonics and the amplitude of these harmonics is defined by Eqn. 1.

We can also generate harmonic stacks in a spectrogram by clipping of the sinusoidal signal at a level ± A_c;

$$y(t) = A \cdot \sin\left(\omega_s \cdot t\right), \qquad -A_c < y < A_c \qquad A_c \in [0, A] \quad \text{Eqn. 2}$$

$$y(t) = A_c, \qquad |y| \geq |A_c| \quad \text{Eqn. 3}$$

We will refer to this case as symmetrical clipping (Figure 1b). See Appendix B2 for the analytical solution of the Fourier integrals. Clipping can also occur asymmetrically, i.e. for only or $-A_c < y$ or $y < A_c$. Here, the boundary A_c is absolute, and in that sense 'hard', in contrast to 'soft', where clipping is implemented with a smooth transition at the clipping level, e.g. with a non-linear term or look-up table. This would also result in a harmonic stack in the spectrogram, but with less energy in the higher frequencies. Even for $A_c = 0.99 \cdot A$, harmonics are still visible in the spectrogram (Figure 1c), which implies that even the slightest clipping of a recorded sound signal may result in false harmonics in the spectrogram. While symmetrical clipping results in the occurrence of odd harmonics

(Figure 1b and 1c), asymmetrical clipping results in both odd and even harmonics (Figure 1d).

In this context, periodic pulse trains can be considered as a special case of a periodic signal. Fast series of discrete pulses are often produced by animals. Examples include stridulation in insects (e.g. Gerhardt & Huber 2003) and birds (Bostwick & Prum 2005), glottal pressure pulses (e.g. Titze 2002), echolocation clicks in odontocetes (Au 1993) and bats (e.g. Holland *et al*. 2004). Depending on the pulse-rate and on the time-window used to build the spectrogram, these signals can appear on spectrograms as very harmonically rich sounds. For a discussion on this specific topic see e.g. Watkins (1967) or Bradbury & Vehrencamp (1989).

Modulations

Another ubiquitous phenomenon in animal vocalisation is modulation. Animals may modulate a carrier signal in various ways: amplitude modulation (AM), frequency modulation (FM) and phase modulation (PM). To generate modulations, we extend our analysis to (linear or non-linear) interactions between sine waves.

Amplitude-modulation

The simplest type of modulation is a sinusoidal amplitude modulation of a sinusoidal carrier signal:

$$y(t) = A \cdot \left(1 + m \cdot y_{AM}\right) \cdot \cos\left(\omega_C \cdot t\right) \text{ for } \omega_{AM} < \omega_C, \qquad \text{Eqn. 4}$$

where $y(t)$ is the modulated time signal, m is the modulation-depth of the signal amplitude, $y_{AM}(t) = \cos(\omega_{AM} \cdot (t)$, ω_{AM} is the modulating signal, is the modulation frequency and ω_c is the carrier-frequency. This equation can be expanded from a product of cosines into a sum of cosines (Spiegel 1974):

$$\begin{aligned} y(t) &= A \cdot \cos\left(\omega_C \cdot t\right) + A \cdot m \cdot \cos\left(\omega_{AM} \cdot t\right) \cdot \cos\left(\omega_C \cdot t\right) = \\ &= A \cdot \cos\left(\omega_C \cdot t\right) + \frac{A \cdot m}{2} \cdot \left[\cos\left(\{\omega_C + \omega_{AM}\} \cdot t\right) + \cos\left(\{\omega_C - \omega_{AM}\} \cdot t\right)\right] \end{aligned} \qquad \text{Eqn. 5}$$

Three frequencies emerge: the carrier-frequency and the sum- and difference-frequencies of carrier- and modulation signal (Figure 2a). Note that the modulation-frequency (ω_{AM}) itself is not visible in the spectrum. The modulation signal (and a component at $\omega = 0$) appears when a DC offset is added to the carrier signal in Eqn. 5 (Bradbury & Vehrencamp 1998; Lavenex 1999; Nowicki & Capricana

1986a, 1986b). In animal sounds, the frequency of the modulation signal is often very low compared with the carrier-frequency (Stein 1968; Watkins 1967). Then, the sum- and difference-frequencies are invisible in the spectrogram and amplitude modulation is only visible in the time signal.

Both carrier and modulation signal can be either sinusoidal or non-sinusoidal, resulting in four possible principle interactions (Figure 2a-d). The carrier signal used in Figure 2 was 3 kHz, and the modulation signal 400 Hz. Non-linearity was introduced to either the modulating or carrier signal, or both, by soft clipping. Despite minute changes in the time domain, the spectra are very different. If the modulation signal is non-sinusoidal and the carrier signal is sinusoidal (Figure 2b), the frequency spectrum above the carrier frequency is an in amplitude scaled duplicate of the spectrum of the modulating signal. The spectrum below the carrier is the mirror image. With a non-sinusoidal carrier signal and a sinusoidal modulation signal, sum and difference frequencies appear around the individual bands of the harmonic stack of the carrier signal (Figure 2c). When both signals are non-sinusoidal, many frequency bands appear around every band of the harmonic stack (Figure 2d). In an engineering context this phenomenon is called intermodulation.

Another type of amplitude modulation is obtained by removing the carrier-frequency from Eqn. 5:

$$y(t) = A \cdot m \cdot \cos\left(\omega_{AM} \cdot t\right) \cdot \cos\left(\omega_C \cdot t\right) = \frac{A \cdot m}{2} \cdot \left[\cos\left(\{\omega_C + \omega_{AM}\} \cdot t\right) + \cos\left(\{\omega_C - \omega_{AM}\} \cdot t\right)\right] \qquad \text{Eqn. 6}$$

This type of modulation is known as Double Side Band modulation (DSB; see Appendix B3) or Amplitude Modulation with suppressed carrier (Figure 2e). It is obtained for any product of sine- and cosine functions by applying simple goniometric rules (Spiegel 1974). It can also be interpreted as the sum of two sinusoidal signals of equal amplitude, but with different frequencies. In the two types of amplitude modulation given by Eqn. 5 and Eqn. 6, the modulation-signal is not present in the spectrum of the modulated carrier-signal. It can be restored by demodulation by a non-linear network as used in AM radios. In Single Side Band modulation (SSB), subsequently, also one of the sideband frequencies $\omega_c - \omega_{AM}$ or $\omega_c + \omega_{AM}$ is removed (Figure 2f) from Eqn. 6. Because the carrier does not contain any additional information, only one sideband is transmitted and radio receivers for SSB need to regenerate the carrier frequency ω_c before demodulation. Although this type of AM might not seem directly relevant to biological systems, it is important to realize that each

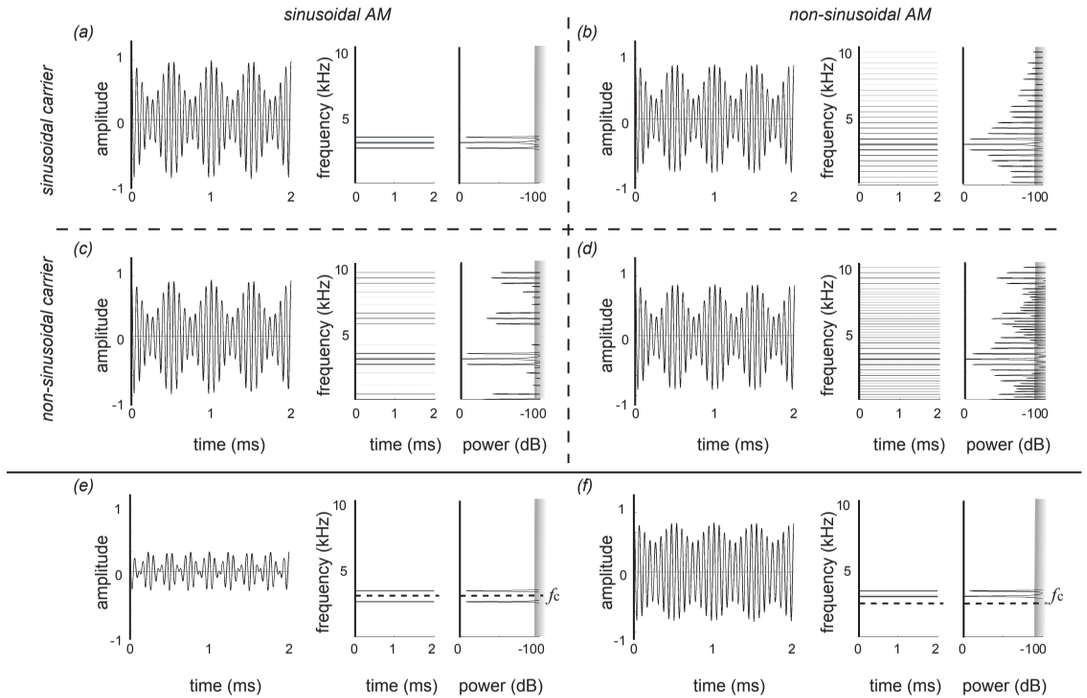

Figure 2. Effects of amplitude modulation (AM). *(a)-(d)* four possible interactions with different carrier and modulation waveforms. *(a)* Simplest form of AM as described in Eqn. 5 resulting in the carrier-frequency and two side-frequencies. *(b)* Non sinusoidal modulation frequency or, *(c)* carrier frequency. *(d)* Intermodulation. *(e)* Double Side Band modulation (Eqn. 6). The carrier-frequency has disappeared, here only the side-frequencies emerge. *(f)* Single Side Band modulation. The carrier-frequency and one side-frequency remain. The latter example can be varied in different ways. ω_c = 3 kHz; ω_{AM} = 400 Hz. F_s = 40 kHz; *nfft* = 1024.

sideband on its own contains all the information of an AM signal. The characteristic amplitude envelope in Figure 2f is often referred to as 'beating' (e.g. Lavenex 1990; Bradbury & Vehrencamp 1998). This spectrum containing two fixed components can be the result of either addition of two sinusoidal signals or the application of Double Side Band modulation.

Summarising, the above formulae show the following. First, amplitude modulation always requires interaction between two or more signals. Second, AM is obtained by adding up signals as well as multiplication. In the context of mathematical operation, the terms 'linear' and 'nonlinear' for respective addition and multiplication sinusoidal signals are indiscriminate and inappropriate. Multiplication still results in a linear system, but not in a time-invariant system. Third, the different types of AM generate very different and intricate spectra, while it is difficult to distinguish between the signals in the time domain.

Angular-modulation

Instead of the two side-frequencies that emerge from sinusoidal amplitude modulation, frequency-modulation gives rise to multiple sidebands of frequencies. At low modulation frequencies these sidebands are not visible in the spectrogram (Figure 3a), but at high modulation frequencies the differences between carrier and FM signal are high enough to visualise the bands (Figure 3b) depending on the bandwidth settings of the spectrogram. Frequency Modulation (FM) and Phase Modulation (PM) are considered to be particular forms of so-called angular modulation (Ziemer & Tranter 1990). Angular modulation is defined as:

$$y(t) = A\sin\left(\omega_t t + \varphi_t\right), \qquad \text{Eqn. 7}$$

where φ_t is a phase constant. In this context, sinusoidal frequency modulation is described by using $\omega_t = \omega_c + m\ cos\ pt$, and $\varphi_t = 0$ in Eqn. 7, where, ω_c is the angular frequency of the carrier and p is the angular frequency of the modulation. For phase modulation with sinusoidal modulation we write $\omega_t = \omega_c$, and $\varphi_t = \varphi_{max} \cos pt$ in Eqn. 7, where φ_{max} is the peak value of the phase deviation. The expression for the instantaneous angular frequency $\omega = d\varphi/dt$, relates FM to PM. The ratio between angular frequencies of the carrier over the modulation, m/p, is the so-called modulation index of a FM signal. Interestingly, the relation of angular modulation (Eqn. 7) can be expanded into underlying Bessel functions that describe the amplitude of each of the individual sidebands created (See Appendix B4). Because a sideband vanishes completely at particular values of m/p, Bessel functions can be used for precise calibration of the modulation process if wanted. In many vocalisations, the modulation

increases in time, which is called progressive FM (Appendix B5) and can be modelled by various means; e.g. either the carrier frequency (Figure 3c) or the modulation index (Figure 3d) can be varied in time.

Generation of multiple signals

With the mathematical principles that were derived above, signals can be generated that contain harmonics, side bands and frequency modulations of an exactly known mathematical origin. Figure 3e shows an example of a synthetically generated signal with combinations of Eqns. 5, 8 and B20 (Appendix B5). Visual inspection of the spectrogram and careful listening to the sound reveals the separate signals. Visual inspection of the time signal, however, gives the impression of aperiodicity.

Nonlinear dynamics

Many sounds in vertebrate animals are produced by structures that can be modelled as nonlinear oscillators. In biological systems a profound source of non-linearity is the non-Hookian relation between applied force and elongation in elastic tissue. These nonlinear oscillators exhibit specific dynamics of which some signatures can be observed in spectrograms (Fee *et al.* 1998; Fitch *et al.* 2001; Fletcher 2000; Wilden *et al.* 1998; Laje & Mindlin 2005). Examples are sudden frequency jumps or doublings, or transitions from harmonic stacks to subharmonics to deterministic chaos (that could mistakenly be considered as noise) in spectrograms. These signatures in spectrograms of animal vocalisations are increasingly reported (e.g. Wilden *et al.* 1998; Laje & Mindlin 2005; Suthers *et al.* 2006). It is beyond the scope of this paper to describe the structures in spectrogram that can be identified as fingerprints of nonlinear dynamics and we like to refer to Wilden *et al.* (1998), Fitch *et al.* (2001) and Laje & Mindlin (2005) for comprehensive descriptions.

However, we would like to point out that nonlinear oscillators exhibit a conversion from AM to PM when the coefficients in their describing differential equations are excitation dependent (Kharkevitch 1962). We will illustrate this AM to PM conversion for an electric circuit, i.e. a nonlinear-RLC series circuit that is voltage excited. The detailed mathematical description of this system can be found in Appendix B6. To show the AM to PM conversion of this system, we concentrated only on the terms concerning the fundamental frequency ω_0 (Appendix B6). Figure 4a shows the amplitude of the charge as a function of normalised frequency. When the excitation increases, the

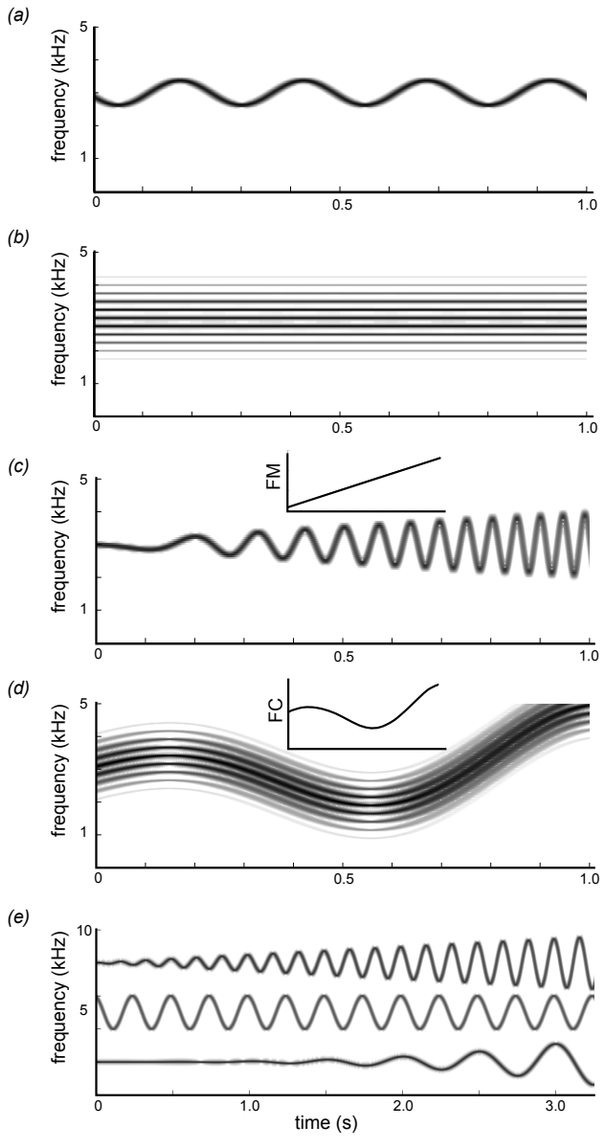

Figure 3. Effects of frequency modulation (FM). *(a)* Normal FM as described in Eqn. 7, with ω_c = 3 kHz and with p = 4 Hz and *(b)* with = ω_{FM} 250 Hz. This results in a spectrum with a carrier-frequency and multiple sidebands of frequencies. *(c)* Progressive FM with varying p and *(d)* with varying ω_c. *(e)* Multiple signals as described in Appendix B5, Eqn. B21.

frequency clearly shifts to a higher value. The phase of the charge oscillation is also strongly dependent on the excitation amplitude (Figure 4b). The AM to PM conversion is evident even in the simplest of all nonlinear differential equation. For large drive amplitudes, these systems can also exhibit the transition from a deterministic to a chaotic oscillator, in the case of which any correspondence between the excitation and the resulting spectra is lost.

From spectrogram to the origin of signals

We demonstrated that many features that are found in the spectrograms of animal vocalisations can be generated with permutations and interactions of simple sine waves. From this approach, we can conclude that signals, which look very similar in the time domain can result in very different spectrograms (e.g. Figure 2). However, what is the explanatory power of the inverse procedure? What can we say about potential underlying mathematical or even mechanical events when looking at the spectrogram? We will demonstrate the diagnostic value of spectrogram analysis by two examples, the vocalisation of a bird, and sound produced by a simple mechanical model.

Example 1. Simulating animal sounds: the screech of a bearded vulture

Birds possibly harbour the largest variation of vocalisations found within the animal kingdom. Therefore most features as described

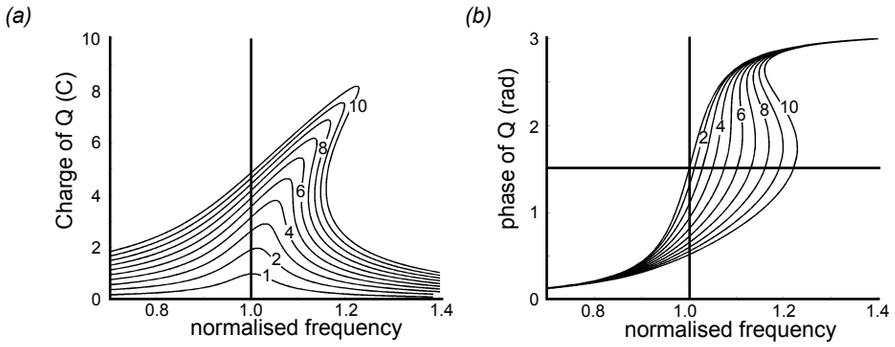

Figure 4. Amplitude modulation transitions to phase modulation in a simple nonlinear oscillator. *(a)* In a simple nonlinear model (see Appendix B6 for a detailed description), a higher excitation amplitude (amplitude of charge *Q*) results in a frequency shift. *(b)* Also the phase shifts as a function of ω^2/ω_0; B = 1, 2, 3, …10, λ = 0.01 and d = 0.1 in Eqns. B28 and B29. For a detailed explanation see Appendix B6.

above can be found in the vocalisations of birds. For this example, we chose the screech of a Bearded Vulture *Gypaetus barbatus* L. The time signal (obtained from *www.birding.dk*) and its corresponding spectrogram are shown in Figure 5a and 5c. This recording was chosen for two reasons: 1) the sound signal contains interesting features, and 2) we do not have access to any physiological or morphological data of the species, because it is highly endangered. Therefore, we only have the data available in the sound signal to deduce its mathematical and mechanical origin.

As can be seen in Figure 5, many energy bands are present that converge towards one band at 1000 ms. In the middle of box a (Figure 5c), bands are visible at 2400, 3100, 3700, 4350, 5000 and 5650 Hz, separated by about 650 Hz. If this would be part of a harmonic stack with a fundamental of 650 Hz, we expect that the harmonics would be spaced at integer multiples of 650Hz (i.e. 1300, 1950, 2600, 3250, 3900, 4550, 5200 and 5850 Hz, respectively). These expected multiples (see Eqn. 1), however, do not coincide with the observed frequencies. Alternatively, the signal could be caused by amplitude modulation, which would require two or more sidebands around a carrier frequency. AM with a carrier frequency of 4550 Hz (or 3900 Hz) and a non-sinusoidal modulation frequency of 650 Hz, generates side bands at the observed frequencies. Also the convergence of the sidebands towards one (carrier) frequency (Figure 5c box b) makes amplitude modulation with a non-sinusoidal modulation signal a very plausible hypothesis. It this case, the bands in the spectrogram are sidebands, (technically also harmonics of the two sidebands in true AM) with one being the carrier frequency. In part (2) and (3) of the spectrogram, we observe a carrier that is frequency-modulated from 3 to 1.5 kHz with some harmonic contents at odd and even multiples. Particularly fast modulations are indicated by arrows on the x-axis in Figure 5c. We made the distinction between part (2) and (3) due to the amplitude modulation visible in the time domain.

Reconstructing this complex signal helps in making predictions about the underlying mathematical events. Furthermore, it provides parameters for statistical analysis. We used the observations noted above to synthesise the vulture's signal. The results of the simulation are shown in Figure 5b and 5d. The first part (1) is a non-sinusoidal AM of a sinusoidal carrier. The carrier signal was constructed with combinations of exponential power functions and the modulation signal was an asymmetrically clipped sine wave. The resulting time signal was clipped asymmetrically to obtain the harmonics in part (2) and (3) of the signal. The AM also results in the sum and difference frequencies in part (1) (see arrow in Figure 5c). This procedure can be extended and improved until the desired match with the original signal is made. Listening to these sounds revealed a close match.

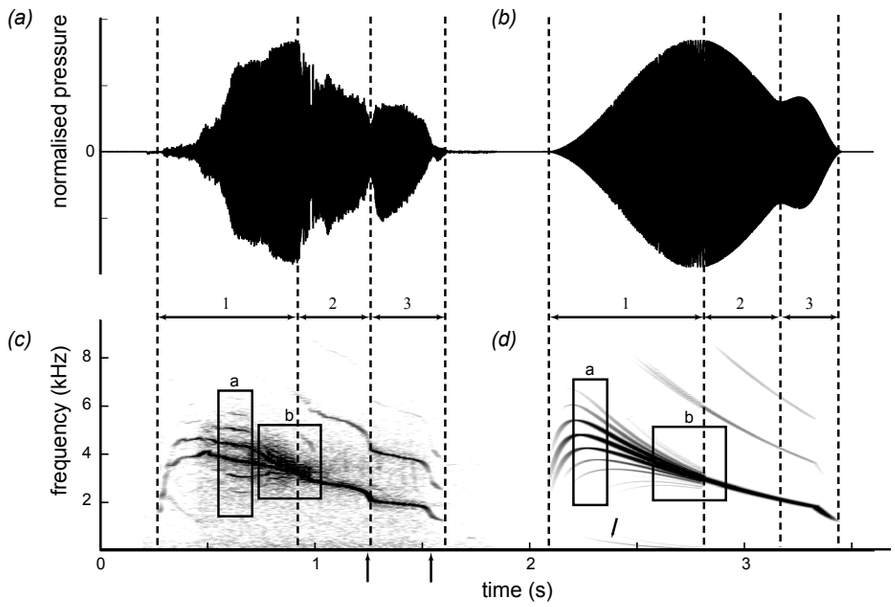

Figure 5. Simulation of a Bearded Vulture *Gypaetus barbatus* screech. The left panels of the figure show the (*a*) time trace and *(c)* spectrogram of an original recording. The right panels of the figure show the *(b)* time trace and *(d)* spectrogram of the simulated signal. To simulate the original sound file, we used the perturbations and modulation types described in this paper. See text for further explanation.

Because we were able to reproduce the original screech using simple mathematical perturbations, we have a strong indication what kind of mathematical phenomena underlie the production of the sound. Furthermore, this approach provided us direct statistics that are useful if one were interested in a comparative approach. What does the suggestion that this sound was generated some form of interaction between two signals – causing the amplitude modulation – tell us about the mechanical origin of the sound? At this point, the explanatory power of solely the spectrogram diminishes considerably. We now need to look at the knowledge we gained in phonation physiology.

The bearded vulture belongs to the order Falconiformes (Sibley & Monroe 1990), and is considered a non-songbird. In non-songbirds, amplitude modulation is reported e.g. in parts of the contact calls and in mimicked English vowel sounds (Lavenex 1999) in Budgerigars *Melopsittacus undulatus*. In general, non-songbirds have a syrinx with only one syringeal aperture (King 1968) formed by two tympaniform membranes, except oilbirds, nightjars and some tropical swifts that have two bronchial apertures (King 1968; Suthers & Hector 1982;

Suthers & Hector 1985). Vibrating tympaniform membranes are the most important sound generator in two non-songbirds; the pigeon *Columba livia* and the cockatiel *Nymphicus hollandicus* (Goller & Larsen 1997; Larsen & Goller 1999). In the light of these findings, we can hypothesise that also in the bearded vulture the tympaniform membranes generate the sound. Based on our reconstruction, AM is the most likely interaction, therefore we need two signals. What mechanical mechanism can cause the AM?

Interference by nonlinearity in the recording equipment of two or more simultaneously vocalising individuals in recordings can cause sidebands (Frommolt 1999), but let us assume that the AM is generated without these types of interference. One possibility is that two separate mechanisms generate the carrier and the modulation signal, i.e. the tympaniform membranes generated the carrier frequency that was modulated by a second mechanism or structure. In the ring dove *Streptopelia risoria*, extremely fast syringeal muscles modulate the sound by pulling the syringeal membranes (i.e. lateral tympaniform membranes or LTM) out of the tracheal lumen (Elemans *et al*. 2004, 2006). The fastest known vertebrate striated muscles, however, contract up to 270 Hz (Fine *et al*. 2001) with a low power output as a trade-off for speed (Rome *et al*. 2006). In both the Budgerigar and the Bearded Vulture, however, modulation frequencies are in the order of 100-700 Hz (Lavenex 1999; Stein 1968), which is too high for any known muscular contraction in vertebrates. Another possible explanation is that more downstream structures interact with the carrier produced in the syrinx, such as flow induced laryngeal movement.

Some types of oscillators provide both the carrier and the modulation, which makes the separation of carrier and modulation signal into different mechanisms not appropriate. For instance, some electronic oscillators have a linear relationship between amplitude and supply voltage, which makes them good amplitude modulators. In the non-oscine syrinx, AM could be generated for instance by the flow or pressure that excites the syringeal membranes.

Another explanation could be that the two tympaniform membranes vibrate at different frequencies. This phenomenon is called 'biphonation' and has been studied in various normal human vocalisations and pathologies (Mergell & Herzel 1997; Tigges *et al*. 1997). It is mainly due to some asymmetry in the elastic tissue properties of the membranes and the amount of coupling between the vocal folds and the vocal tract (Mergell & Herzel 1997). So, although biphonation occurs only under specific conditions, it is plausible that the physiology and morphology of different species has evolved to meet them.

All these hypotheses are not easily tested for our screeching vulture. In songbirds, amplitude modulation has previously been reported in the Black-capped Chickadee *Parus atricapillus* by

Nowicki & Capricana (1986a, 1986b). This species belongs to the Order Passeriformes (the songbirds) that have two, bilateral syringeal apertures. By denervation of both sides of the syrinx, Nowicki & Capricana (1986a, 1986b) provided evidence that both apertures produced a carrier with a different frequency. AM was the result of coupling between the left and right parts of the syrinx. To find the origin of the AM in our screeching vulture, we need detailed information on the vibratory behaviour of the sound generators. Obtaining direct measurements of syringeal vibration is obviously very challenging, due to the technical difficulties involved.

Concluding, the answer to our initial question if we can derive the mechanical origin of the sound from our spectrogram is negative. However, we can postulate multiple hypotheses about which potential mathematical processes are underlying the signal, and test these by comparing synthesised features with the observed features in the spectrogram.

Example 2. The use of a mechanical model, the bagpipe reed

In our second example, we employ a mechanical model to produce sound. In contrast to *in vivo* measurements of sound generators in biological systems, mechanical models allow for much more accurate measurements of the mechanical events at relative low effort. Furthermore, mechanical models serve to test the accuracy of mathematical models, and certain parameters can be changed that are fixed in the biological system, such as material properties. Mechanical models are ubiquitously used in engineering and quite common in the fields of biomechanics (e.g. Koehl 2003) and human voice modelling (de Vries 2003).

A simple mechanical model that generates sound is a drone reed from a bagpipe consisting of a cane tube *Arundo donax* L.. The cane tube is closed on one side and transversally and longitudinally cut to form a blade (Figure 6a). Abduction of the blade widens the slit in the tube (Figure 6a and 6b). The player does not manipulate the blade directly with their mouth such is the case in a clarinet or saxophone. The drone is connected to the chanter, of which the resonance properties are determined by the player by closing and opening holes. The blade of the reed and the air in the reed tube are coupled oscillators. The design of the blade and its exact geometry has been subject to many hundreds of years of cultural evolution (Baines 1973). In a bagpipe, a pressure load is applied via the bag to induce vibrations of the reed. We mounted the reed in the sidewall of an acrylate cylinder (inner diameter 34 mm) to mimic the bag and to allow for optical access. The cane tube was open on

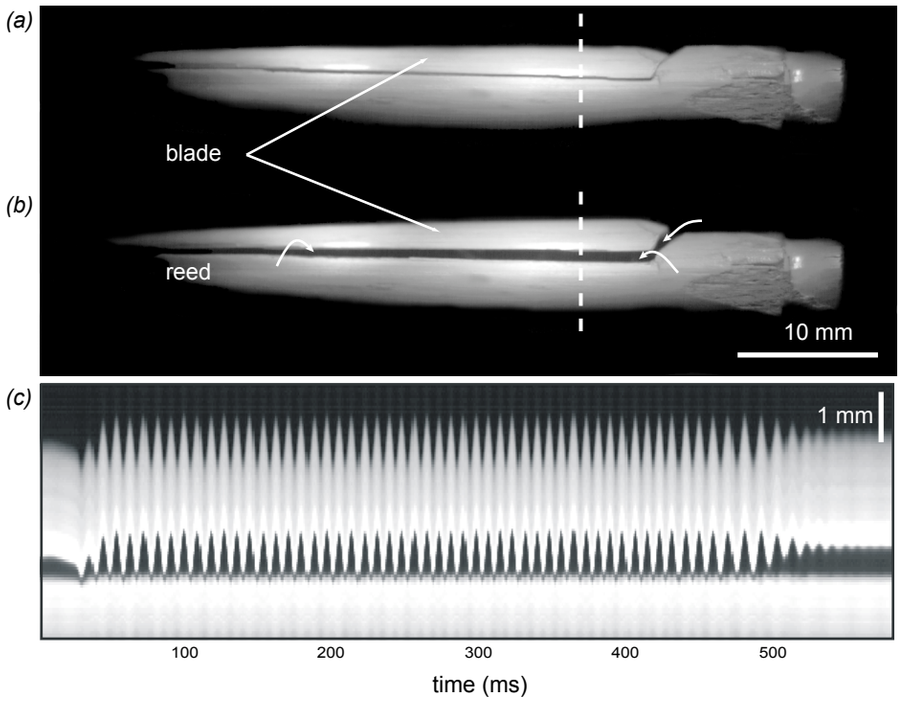

Figure 6. Close-up of the mechanical model used, the bag-pipe drone. The two top images are from the high-speed recordings and show the position of the blade when fully closed *(a)* and open *(b)*. The arrows indicate the air flow that runs from the right through the tube towards the outlet (on the left, not shown). *(c)* The intensity of the drone model images at a fixed position (dotted line in *(a)* and *(b)*) plotted as a function of time, nicely shows the movement of the blade in time. The oscillation period equals 4 ms.

one side and formed an outlet. The pressure in the cylinder could be regulated with pressurised air. Sound was recorded with a condenser microphone (Brüel and Kjær 4939; Preamplifier 2670) at 30 cm from the tube-outlet of the reed in a semi-anechoic room at Wageningen University. The signal was amplified and band-pass filtered (20Hz (40 dB/dec) –10 kHz (40 dB/dec), Nexus amplifier; Brüel and Kjær). Gauge pressure was measured about 10 mm from the end of the blade inside the tube with a cathetertip pressure transducer (Millar SPC 350, diameter of tip 1.6mm) and low-pass filtered (10kHz) with a custom-built analogue filter. Pressures were measured relative to the ambient pressure. All signals were digitised at 30 kHz using a 12-bit AD-converter (model PCI-MIO16E-4, National Instruments) on a Pentium III 700MHz Workstation. After digitising, the pressure signal was band-pass filtered (100 Hz-10 kHz) with a 3rd order Butterworth digital filter using zero-phase implementation (Oppenheim & Schafer

1989). Monochrome, digital high-speed video recordings of the moving blade were made at 2 kHz (Redlake, MotionPro). The video system was synchronised with the sound and pressure signals with a time difference less than 0.5 ms. The width of the opening between the blade and the reed was measured from the high-speed images at a fixed position (dotted line in Figure 6). Width opening as a function of time was up-sampled from 2 to 30 kHz by cubic spline-interpolation (de Boor 1987; Press *et al.* 1990) through the time data. Signal acquisition and analysis software was developed in Matlab (The Mathworks).

Time traces and spectrograms of the pressures inside and outside the tube of a typical recording are shown in Figure 7. The sound pressure trace varied in amplitude from 12 to -9 Pa (Figure 7a). The pressure in the tube varied from about 80 to -350 Pa (Figure 7b) around atmospheric pressure. Both the sound pressure and the pressure in the tube exhibited a consistently repeating pattern with a period of 4 ms (Figure 7c, d). Since both pressure functions were non-sinusoidal but repetitive, it was likely that odd and even harmonics of the 250 Hz fundamental will occur in their spectra as long as the time window for the FFT is sufficiently long to capture more than a single period.

Calculated spectrograms (Figure 7e, f) consisted indeed of a harmonic stack with a fundamental of 250 Hz and harmonics at evenly spaced odd and even integer multiples (500, 750, 1000, … Hz). The harmonic stack was found in both spectrograms, but their spectra were different since the waveforms differed in shape and amplitude.

What can we conclude from these signals using the mathematical theory presented above? The simplest way to generate a harmonic stack structure is by either asymmetrical clipping or full wave rectifying as described in Section 2.2. A more complex way to generate the structure in the spectrogram is by non-sinusoidal AM or FM with application of specific filtering. For example, AM with a sinusoidal carrier of 3 kHz and a non-sinusoidal modulation signal with a fundamental of 250 Hz generates multiple side-bands as shown in Figure 5B. In this case, the sidebands overlap exactly (since 3000/250=12, an integer) and could generate a structure similar to Figure 7e. Additionally, a specific filter could match the spectra. The hypothesis that the signal is generated with AM, however, can be rejected because the bands in the spectrogram remain harmonically related also when the lowest band shows slight frequency modulation at the last 400ms. The harmonic relation during FM demonstrates that the most parsimonious explanation is that the lowest band is the true fundamental frequency of the harmonic stack.

In Section 2.2, we showed that a harmonic structure could be generated if somewhere in the system there is a form of clipping, or more broadly defined, a unilateral restriction. The structure in

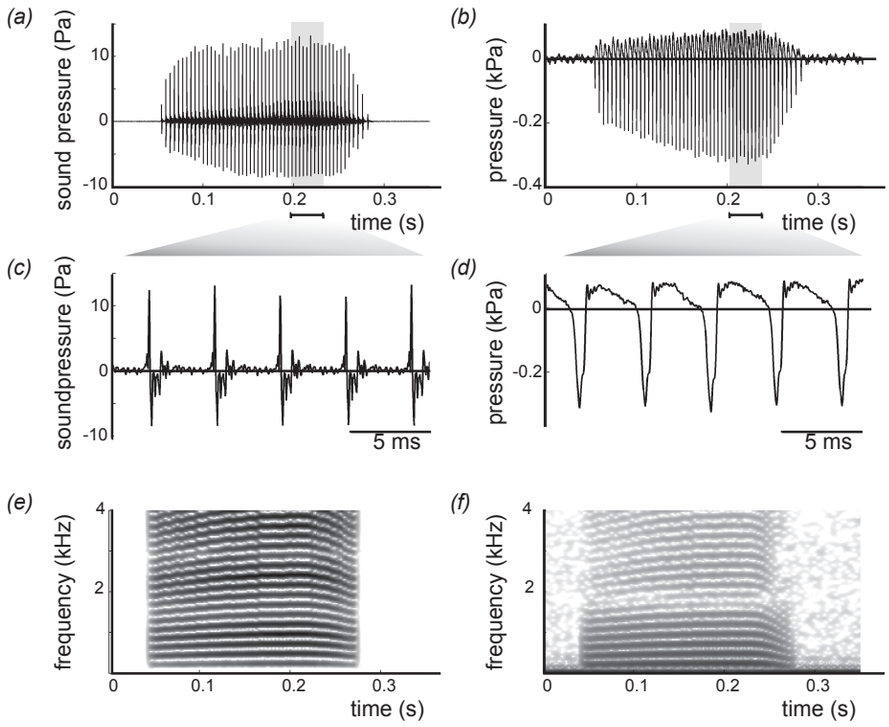

Figure 7. Pressure recordings in and outside of the bag-pipe drone. *(a)* Sound pressure at 30 cm from the tube outlet and *(b)* gauge pressure measured in the tube (relative to atmospheric pressure). *(c)-(d)* Details of *(a)* and *(b)*. *(e)-(f)* Spectrograms of the signals in *(a)* and *(b)*. F_s = 30kHz; *nfft*= 1024; df = 29.4 Hz; dt = 17ms. The strong harmonic stacks in both spectrograms are caused by the pulsatile, but repetitive signals. For more explanation see text.

the spectrogram could indicate a unilateral restriction in the sound generator. Note that even when the signal was generated by AM of some sort, it would require a non-sinusoidal modulation frequency and thus some non-linearity.

During sound production, the blade periodically narrows (Figure 6a) and widens (Figure 6b) the slit. The motion of the blade as a function of time during one bout is visualised in Figure 6c. The high-speed video recordings showed that the blade vibrated in its basal mode and was clearly restricted in its movement by the tube (Figure 8a). Since the pressure in the acrylate cylinder is above ambient room pressure, the air flowed (indicated by arrows in Figure 6a, b) from the acrylate chamber, through the cane tube into the room. During collision between the blade and the tube, the pressure in the tube decreased from 0 Pa to -300 Pa and then increased sharply (Figure 7d). When the blade closed the passage of air, the air in the tube moved on

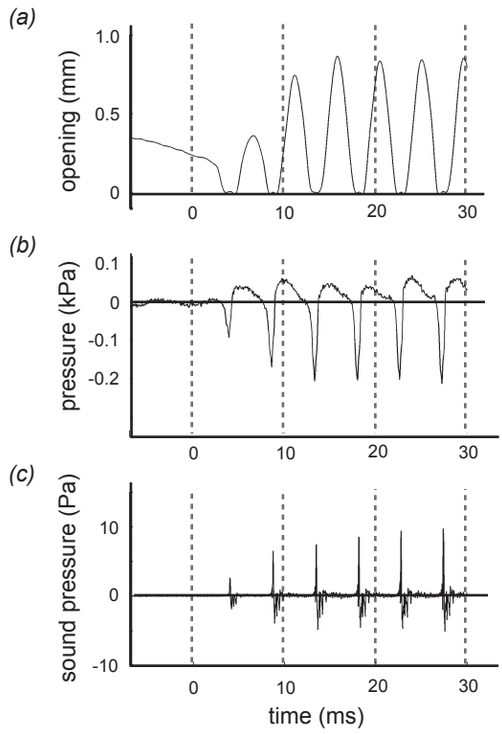

Figure 8. Detail of onset of sound production in the drone model. *(a)* Amplitude of the reed opening. *(b)* Pressure in the tube and *(c)* sound pressure at 30 cm from the tube outlet.

due to its momentum. The local density of the air close to the blade decreased and pressure dropped below ambient pressure. Inertia of the air caused this rapid decrease in pressure. Subsequently, air was sucked through the outlet of the tube, which increased pressure until the blade was pushed upwards. At this stage, air flowed again through the opening of the blade. The movement of the blade is determined by intrinsic mechanical properties and geometry. The movement of the blade and the other mechanical events are locked in periodicity and display limit cycle behaviour comparable to the vocal folds.

Summarising, the analysis of the structure in the spectrogram again helped to hypothesise about the clipping event and thus directed the search for identification of the sound source. The spectrogram, however, failed to explain the mechanical underlying processes of the measured signals. The origin of the harmonics in the spectrogram is the periodicity of the waveforms. Understanding the measured signals required much more knowledge of the flow dynamics and aero-acoustics of the system. (see e.g. Deverge *et al.* 2003; Fabre & Hirschberg 2000; Verge *et al.* 1997a, 1997b).

DISCUSSION

We started by relating simple structures in spectrograms to their causal mathematical events. We identified immediate signatures of repetition, modulations, transients and nonlinear behaviour. As such, a spectrogram of a vocalisation provides many clues about the mathematical nature of the signal. A spectrogram allows us to define clear hypotheses on how the signals might have been produced based on these clues. However, to interpret these mathematical permutations in the light of meaningful physical events, additional knowledge of sound generation is necessary. About the role of morphology in sound generation, Greenewalt (1968) stated: 'Unfortunately it is not easy to make the transition from an anatomical structure, no matter how detailed, to a description of its functional operation.' From the viewpoint of acoustic analysis, Gaunt & Gaunt (1985) made the remark that 'If (avian) voice is modulated even in part by nonsyringeal mechanisms, then syringeal function cannot be deduced from acoustic analysis of the call alone, regardless of the sophistication of the analysis. Acoustic evidence alone may fit contradictory models equally well.' The spectrogram provides important clues about the origin of animal sounds, but as with morphological studies and studies of system dynamics, detailed studies of the spectrogram provide only a partial insight in the sound producing system.

By looking at our two examples, we can conclude that a spectrogram *per se* leaves us clueless about the underlying mechanical origin of the sound signal. This is especially the case when dealing with nonlinear systems (see Section 2.4). Systems already exist that can reproduce any sound signal which in a spectrogram would be indistinguishable from the original signal. The loudspeaker consists of only one vibrating membrane and is capable of reproducing sound signals from wind noise to bird song, from conversations to orchestras. Animals, however, do not use speakers to generate sound. Although apparently a trivial observation, it is the state of the art knowledge of sound producing mechanisms in animals that conveniently reduces the number of possible explanations. In bioacoustics there are the famous examples where two mechanically different sound production systems can produce very similar sounds; a man who can imitate a steam engine, or the Superb Lyrebird *Menura novaehollandiae* that can imitate a chain saw. Knowledge of sound producing systems does increase the explanatory power of a spectrogram, but a spectrogram per se unfortunately cannot present unambiguous evidence.

ACKNOWLEDGEMENTS

The authors thank Johan van Leeuwen, Simon Bœl Pederson, Stuart

Rosen and anonymous reviewers for comments on earlier versions of the manuscript. We thank Mr. Bjarne Nielsen for providing bird vocalisations on www.birding.dk. The investigations were (in part) supported by the Research council for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organisation for Scientific Research (NWO).

REFERENCES

Abramowitz, M. & Stegun, I. A. (Eds.). (1970). *Handbook of Mathematical Functions*. New York: Dover Publications.

Askew, G. N. & Ellerby, D. J. (2007). The mechanical power requirements of avian flight. *Biology Letters* **3**, 445-448.

Au, W. W. L. (1993). *The Sonar of Dolphins*. New York: Springer.

Baines, A. (1973). *Bagpipes*. Occasional papers on Technology, 9. Edited by T. K. Penniman and B. M. Blackwood. Oxford University Press.

Beecher, M. D. (1988). Spectrographic analysis of animal vocalizations: implications of the "uncertainty principle". *Bioacoustics* **1**, 187-208.

Bradbury, J. W. & Vehrencamp, S. L. (1998). *Principles of Animal Communication*. Canada: Sinauer Ass. Inc.

de Boor, C. A. (1978). *Practical Guide to Splines*. Springer-Verlag.

Bostwick, K.S. & Prum, R.O. (2005). Courting Bird Sings with Stridulating Wing Feathers. *Science* **309**, 736.

Clark, C. W., Marler, P. & Beeman, K. (1987). Quantitative analysis of animal phonology: an application to swamp sparrow song. *Ethology* **76**, 101-115.

Cohen, L. (1995). *Time-frequency analysis: Theory and applications*. Prentice-Hall Signal Processing series.

Cortopassi, K. A. & Bradbury, J. W. (2000). The comparison of harmonically rich sounds using spectrographic cross-correlation and principal coordinates analysis. *Bioacoustics* **11**, 89-127.

Darden, S. K., Pedersen, S. B. & Dabelsteen, T. (2003). Methods of frequency analysis of a complex mammalian vocalisation. *Bioacoustics* **13**, 247-263.

Deverge, M., Pelorson, X., Vilain, C., Lagrée, P.Y., Chentouf, F., Willems, J. & Hirschberg, A. (2003). Influence of the collision on the flow through in-vitro rigid models of the vocal folds. *J. Acoust. Soc. Am.* **114**, 3354-3362.

Duffing, G. (1918). Erzwungene Schwingungen bei veränderlicher Eigenfrequenz. Vieweg. Braunschweig.

Elemans, C. P. H., Spierts, I. L. Y., Muller, U. K., van Leeuwen, J. L. & Goller, F. (2004). Superfast muscles control dove's trill. *Nature* **431**, 146.

Elemans, C. P. H., Spierts, I. L. Y., Hendriks, M., Schipper, H., Müller, U.K. & van Leeuwen, J. L. (2006). Syringeal muscles fit the trill in ring doves (*Streptopelia risoria* L.) *J. Exp. Biol.* **209**, 965-977.

Fabre, B. & Hirschberg, A. (2000). Physical modelling of Flue instruments: A review of lumped models. *Acta Acustica* **86**, 599-611.

Fee, M. S., Shraiman, B., Pesaran, B. & Mitra, P. P. (1998). The role of nonlinear dynamics of the syrinx in the vocalisations of a songbird. *Nature* **395**, 67-71.

Fitch, W. T., Neubauer, J. & Herzel, H. (2001). Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behaviour* **63**, 407 – 418.

Fine, M. L., Malloy, K. L., King, C. B., Mitchell, S. L. & Cameron, T. M. (2001). Movement and sound generation by the toadfish swimbladder. *J. Comp. Phys. A* **187**, 371-379.

Fletcher, N. H. (2000). A class of chaotic bird calls? *J. Acoust. Soc. Am*. **108**, 821-826.
Frommolt, K.-H. (1999). Sidebands – facts and artefacts. *Bioacoustics* **10**, 219-224.
Gardner, T. J. and Magnasco, M. O. (2005). Instantaneous frequency decomposition: An application to spectrally sparse sounds with fast frequency modulations. *J. Acoust. Soc. Am*. **117**, 2896-2903.
Gardner, T. J. and Magnasco, M. O. (2006). Sparse time-frequency representations. *Proc. Natl. Acad. Sci USA* **103**, 6094-6099.
Gaunt, A. S. & Gaunt, S. L. L. (1985). Syringeal structure and avian phonation. *Curr. Ornith*. **2**, 213-245.
Gerhardt, H. C. & Huber, F. (2003). *Acoustic Communication in Insects and Anurans: Common Problems and Diverse Solutions*. Chicago: University of Chicago Press.
Goller, F. & Larsen, O. N. (1997). *In situ* biomechanics of the syrinx and sound generation in pigeons. *J. Exp. Biol*. **200**, 2165-2176.
Greenewalt, C. H. (1968). *Bird Song: Acoustics and Physiology*. Washington DC: Smithsonian Institute.
Hall-Craggs, J. (1979). Sound spectrographic analysis: suggestions for facilitating auditory imagery. *Condor* **81**, 185-192.
Holland, R. A., Waters, D. A. & Rayner, J. M. V. (2004). Echolocation signal structure in the Megachiropteran bat *Rousettus aegyptiacus* Geoffrey 1810. *J. Exp. Biol*. **207**, 4361-4369.
Jones, D. L. & Barianuk, R. G. (1995). An adaptive optimal-kernel time-frequency representation. *IEEE Trans. Sig. Proc*. **43**, 2361-2371.
Khanna, H., Gaunt, S. L. L. & McCallum, D. A. (1997). Digital spectrographic cross-correlation: test of sensitivity. *Bioacoustics* **7**, 209-234.
Kharkevitch, A. A. (1962). *Nonlinear & Parametric Phenomena in Radio Engineering*. New York: John F. Rider Publisher Inc.
King, A. S. (1989). Functional anatomy of the syrinx, In *Form and Function in Birds. Vol. 4* (Ed. by A. S. King & J. McLelland), pp. 105-192. New York: Academic Press.
Koehl, M. A. R. (2003). Physical modelling in biomechanics. *Phil. Trans. R. Soc. Lond*. B **358**, 1589-1596.
Laje, R. & Mindlin, G. B. (2005). *The physics of bird song*. Berlin: Springer.
Larsen, O. N. & Goller, F. (1999). Role of syringeal vibrations in birds vocalizations. *Proc. R. Soc. Lond. B* **266**, 1609-1615.
Lavenex, P. B. (1999). Vocal production mechanisms in the budgerigar (*Melopsittacus undulatus*): The presence and implications of amplitude modulation. *J. Acoust. Soc. Am*. **106**, 491-505.
Léonard, F. (1997). Referencing the phase to the centre of the spectral window. Why? *Mech. Syst. and Sig. Proc*. **11**, 75-90.
Léonard, F. (2000). Spectrogramme de phase et spectrogram de fréquence. *Traitement du Signal* **17**, 269 – 286.
Léonard, F., Foata, M., & Paquin, J. Y. (2000). Vibro-acoustic signature comparison and time-warping correction with multi-scale correlation. *Mech. Syst. and Sig. Proc*. **14**, 443-458.
Lillesand, T. M. & Kiefer, R. W. (2000). *Remote sensing and image interpretation*. 4th ed., New York: John Wiley.
Mbu Nyamsi, R. G., Aubin, T. & Bremond, J. C. (1994). On the extraction of some time dependent parameters of an acoustic signal by means of the analytical signal concept. Its application to animal sound study. *Bioacoustics* **5**, 187-203.
Mergell, P. & Herzel, H. (1997). Modelling biphonation – The role of the vocal tract. *Speech Comm*. **22**, 141-154.
Nowicki, S. & Capranica, R.R. (1986a). Bilateral syringeal coupling during phonation of a songbird. *J. Neuroscience* **6**, 3595-3610.

Nowicki, S. & Capranica, R. R. (1986b). Bilateral syringeal interaction in vocal production of an oscine bird sound. *Science* **231**, 1297–1299.
Oppenheim, A. V. & Schafer, R. W. (1989). *Discrete-Time Signal Processing*. New York: Prentice-Hall.
Paulsen, K. (1967). *Das Prinzip der Stimmbildung in der Wirbeltierereihe und beim Menschen*. Frankfurt am Main: Akademische Verlagsgesellschaft.
Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1990). *Numerical recipes in Pascal. The art of scientific computing*. Cambridge: Cambridge University Press.
Rome, L. C. (2006). Design and function of superfast muscle: new insights. *Ann Rev Physiol* **68**, 193-221.
Sibley, C. G. & Monroe, B. L. *(1990). Distribution and Taxonomy of Birds of the World*. New Haven and London: Yale University Press.
Spiegel, M. R. (1974). *Theory and problems of Fourier Analysis*. Schaum Outline Series. New York: McGraw-Hill.
Staddon, J. E. R., McGeorge, L. W., Bruce, R. A. & Klein, F. F. (1978). A simple method for the rapid analysis of animal sounds. *Zeit. Tierpsychol.* **48**, 306-330.
Stein, R. C. (1968). Modulations in bird sounds. *Auk* **85**, 229-243.
Suthers, R. A. & Hector, D. H. (1982). Mechanism for the production of echolocating clicks by the grey swiftlet, *Collocalia spodiopygia*. *J. Comp. Physiol. A.* **148**, 457-470
Suthers, R. A. & Hector, D. H. (1985).The physiology of vocalisation by the echolocating oilbird *Steatornis caripensis*. *J. Comp. Physiol. A.* **156**, 243-266.
Suthers, R. A., Narins, P. M., Lin, W. Y., Schnitzler, H. U., Denzinger, A., Xu, V. H. & Feng, A. S. (2006). Voices of the dead: complex nonlinear vocal signals from the larynx of an ultrasonic frog. *J. Exp. Biol.* **209**, 4984-4993
Tchernichovski, O. & Mitra, P. P. (2002). Towards quantification of vocal imitation in the zebra finch. *J. Comp. Phys. A.* **188**, 867-878.
Tigges, M., Mergell, P., Herzel, H., Wittenberg, T. & Eysholdt, U. (1997). Observation and modelling of glottal biphonation. *Acustica* **83**, 707-714.
Titze, I. R. (2002). *Principles of Voice Production*. Englewood Cliffs: Prentice-Hall.
Versluis, M., Schmitz, B., von der Heyft, A. & Lohse, D. (2000). How snapping shrimp snap: Through cavitating bubbles. *Science* **289**, 2114-2117.
Verge, M. P., Fabre, B., Hirschberg, A. & Wijnands, A. P. J. (1997a). Sound production in recorderlike instruments 1. Dimensionless amplitude of the internal acoustic field. *J. Acoust. Soc. Am.* **101**, 2914-2924.
Verge, M. P., Hirschberg, A. & Causse, R. (1997b). Sound production in recorderlike instruments .2. A simulation model. *J. Acoust. Soc. Am.* **101**, 2925-2939.
Von Békésy, G. (1960). *Experiments in hearing*. New York: McGraw Hill.
de Vries, M. P., Hamburg, M. C., Schutte, H. K., Verkerke, G. J. & Veldman, A. E. P. (2003). Numerical simulation of self-sustained oscillation of a voice-producing element based on Navier–Stokes equations and the finite element method. *J. Acoust. Soc. Am.* **113,** 2077-2083.
Wakeling, J. M., Kaya, M., Temple, J. K., Johnston, I. A. & Herzog, W. (2002). Determining patterns of motor recruitment during locomotion. *J. Exp. Biol.* **205**, 359-369.
Watkins, W. A. (1967). The harmonic interval: fact or artefact in spectral analysis of pulse trains. In: *Marine Bio-acoustics Volume 2*. (Ed. by W.N. Tavolga). Oxford: Pergamon Press.
Wilden, I., Herzel, H., Peters, G. & Tembrock, G. (1998). Subharmonics, biphonation, and deterministic chaos in mammal vocalisation. *Bioacoustics* **9**, 171-196.
Williams, J. M. & Slater, P. J. B. (1991). Computer analysis of bird sounds: a guide to current methods. *Bioacoustics* **3**, 121-128.

Ziemer, R. E. & Tranter, W. H. (1990). *Principles of Communications*. Boston, Mass: Houghton Mifflin Co.

Received 2 February 2008, revised 30 June 2008 and accepted 8 July 2008.

APPENDICES

Appendix A: abbreviations and units of the parameters used

A	amplitude [m]	
A_c	clipping amplitude [m]	
AM	Amplitude Modulation	
C	capacitance [F]	
DFFT	Discrete Fourier Transform	
DSB	Double Side Band modulation	
f	frequency [s^{-1}],	$f = 1/T$
FFT	Fast Fourier Transform	
FM	Frequency Modulation	
i	imaginary unit = $\sqrt{(-1)}$	
J_n	Bessel function of n^{th} order.	
L	inductance [H]	
m	modulation depth of amplitude modulation	
n	modulation index of frequency modulation; integer value (n=0,1,2,3...)	
p	angular cycle frequency in angular modulation	
PM	Phase Modulation	
Q	charge [C]	
R	resistance (Ω)	
SBB	Single Side Band modulation	
T	time constant [s]	
t	time [s]	
ω	angular frequency [rad · s^{-1}],	ω = 2 · π · f
ω_C	carrier frequency [rad · s^{-1}]	
ω_{AM}	angular frequency of amplitude modulating signal [rad · s^{-1}]	
ω_{FM}	angular frequency of frequency modulating signal [rad · s^{-1}]	
ω_s	angular frequency of signal s [rad · s^{-1}]	
ω_t	angular frequency in angular modulation [rad · s^{-1}]	

Appendix B: derivations

(B1) Fourier-series of a full wave rectified cosine function

Let $y(t) = A \cdot \cos(\omega_s \cdot t) = A \cdot \sin\left(\omega_s \cdot t + \frac{\pi}{2}\right)$. Eqn. B1

Substitution of the new argument of the above sine-function in the Fourier-series of the full wave rectified sine-function given by Eqn. 1, leads to the Fourier-series of a full wave rectified cosine-function:

$$y(t) = \left|A \cdot \cos(\omega_s \cdot t)\right| = A \cdot \left[\frac{2}{\pi} + \frac{4}{\pi}\left(\frac{\cos(2 \cdot \omega_s \cdot t)}{1 \cdot 3} - \frac{\cos(4 \cdot \omega_s \cdot t)}{3 \cdot 5} + \frac{\cos(6 \cdot \omega_s \cdot t)}{5 \cdot 7} + \ldots\right)\right] \quad \text{Eqn. B2}$$

Note that also here only even harmonics occur and that the fundamental frequency has disappeared.

(B2) Analytical solver for the harmonic contents of a clipped sinusoidal signal

We start from a signal:

$$y(t) = A \cdot \sin(\omega_s \cdot t) = A \cdot \sin\left(\frac{\pi}{L_s} \cdot t\right) \qquad |y| < A_c \qquad A_c \in [0, A] \qquad \text{Eqn. B3}$$

$$y(t) = A_c, \qquad |y| \geq A_c \qquad \text{Eqn. B4}$$

A_c is the level at which the signal is clipped and L_s = ½ T_s. For a half-range expansion of this odd function, it rules: (1) there are only sin-components in the Fourier series, (2) the amplitudes a_n (n = 0,1,2...) in the Fourier series are zero, so only the b_n amplitudes remain:

$$y(t) = \sum_{n=1}^{\infty} b_n \cdot \sin\left(\frac{n \cdot \pi \cdot t}{L_s}\right) \qquad \text{Eqn. B5}$$

The amplitude b_n of this half-range expansion is:

$$b_n = \frac{2}{L_s} \cdot \int_0^{L_s} y(t) \cdot \sin\left(\frac{n \cdot \pi \cdot t}{L_s}\right) dt \qquad \text{Eqn. B6}$$

Using the instants τ_1 and τ_2 at which the clipping level is reached, we can split up this integral in three parts (Fig. B1a):

$$b_n = \frac{2 \cdot A}{L_s} \cdot \int_0^{\tau 1} \sin\left(\frac{\pi \cdot t}{L_s}\right) \cdot \sin\left(\frac{n \cdot \pi \cdot t}{L_s}\right) dt + \frac{2 \cdot A_c}{L_s} \cdot \int_{\tau 1}^{\tau 2} \sin\left(\frac{n \cdot \pi \cdot t}{L_s}\right) dt + \frac{2 \cdot A}{L_s} \cdot \int_{\tau 2}^{L} \sin\left(\frac{\pi \cdot t}{L_s}\right) \cdot \sin\left(\frac{n \cdot \pi \cdot t}{L_s}\right) dt$$
$$= I_1 + I_2 + I_3 \qquad \text{Eqn. B7}$$

in which:

$$\tau_1 = \frac{L_s}{\pi} \cdot \arcsin\left(\frac{A_c}{A}\right) \qquad \tau_2 = L_s - \tau_1 \qquad \text{Eqn. B8}$$

Substituting $y(t)$ and τ in (C8) and carrying out the integration yields for A = 1:

$$I_1 = \frac{1}{\pi \cdot (1-n)} \cdot \sin\left[(1-n) \cdot \arcsin(A_c)\right] - \frac{1}{\pi \cdot (1+n)} \cdot \sin\left[(1+n) \cdot \arcsin(A_c)\right] \qquad \text{Eqn. B9}$$

$$I_2 = \frac{2 \cdot A_c}{n \cdot \pi} \cdot \left[\cos\left(n \cdot \arcsin(A_c)\right) - \cos\left(n \cdot \{\pi - \arcsin(A_c)\}\right)\right] \qquad \text{Eqn. B10}$$

$$I_3 = \frac{1}{\pi \cdot (1-n)} \cdot \left[\sin\{\pi \cdot (1-n)\} - \sin\left[(1-n) \cdot \{\pi - \arcsin(A_c)\}\right]\right] - \frac{1}{\pi \cdot (1+n)} \cdot \left[\sin\{\pi \cdot (1+n)\} - \sin\left[(1+n) \cdot \{\pi - \arcsin(A_c)\}\right]\right] \qquad \text{Eqn. B11}$$

For n = 1, the above formulae contain sinc-functions and become:

$$I_1 = I_3 = \frac{1}{\pi} \cdot \arcsin(A_c) - \frac{1}{2 \cdot \pi} \cdot \sin\left[2 \cdot \arcsin(A_c)\right] \qquad \text{Eqn. B12}$$

$$I_2 = \frac{2 \cdot A_c}{\pi} \cdot \left[\cos\left(\arcsin\left(A_c\right)\right) - \cos\left(\pi - \arcsin\left(A_c\right)\right) \right] \qquad \text{Eqn. B13}$$

The behaviour of these expressions when A_c varies from 0-1 is displayed in Fig. B1**b** and B1c. It follows that even for $A_c/A \rightarrow 1$ the amplitude of the harmonics has still a considerable value. This is understandable as the sine-function about $t = L_s$ is very shallow and therefore the duration of the level A_c remains rather large when $A_c/A \rightarrow 1$.

(B3) Derivation of the formula for Double Side Band modulation (DSB)

Removing the carrier-signal from Eqn. 6 leads to the formula of double sideband modulation:

$$\begin{aligned} y_1(t) &= y(t) - A \cdot \cos\left(\omega_C \cdot t\right) = +A \cdot m \cdot \cos\left(\omega_{AM} \cdot t\right) \cdot \cos\left(\omega_C \cdot t\right) = \\ &= \frac{A \cdot m}{2} \cdot \left[\cos\left(\{\omega_C + \omega_{AM}\} \cdot t\right) + \cos\left(\{\omega_C - \omega_{AM}\} \cdot t\right) \right] \end{aligned} \qquad \text{Eqn. B14}$$

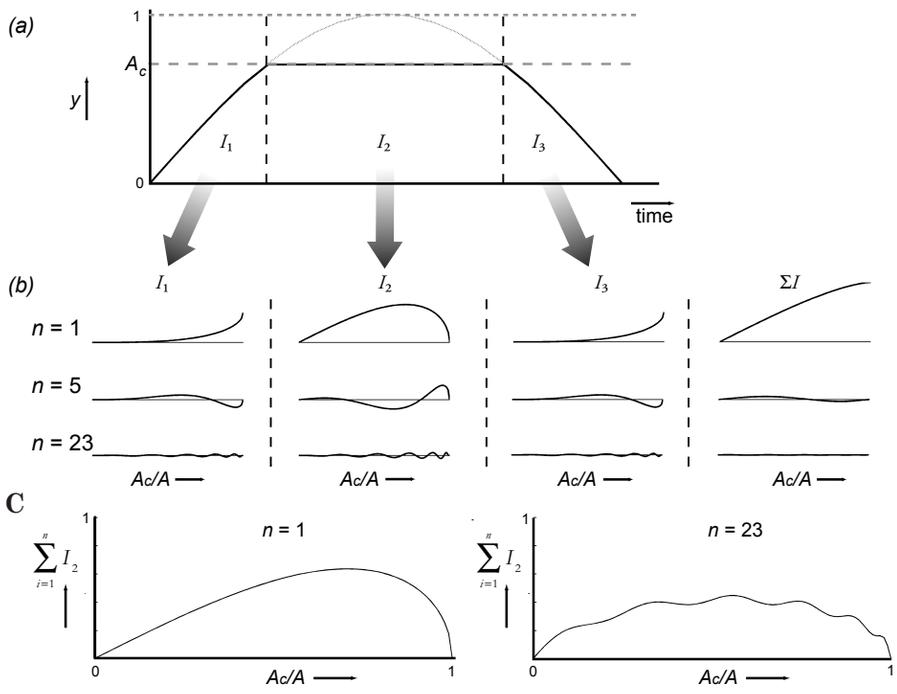

Figure B1. An analytical solver for the harmonic content of a clipped sine wave. *(a)* Half-range expansion of a clipped sinusoidal signal. A = amplitude, A_c = amplitude of clipping-level, τ = clipping-time, L = half-period. *(b)* The Fourier-integrals of the coefficient b_n as a function of A_c for n = 1, 5 and 25. **C** $\sum_{i=1}^{n} I_2$ for n = 1, and $\sum_{i=1}^{n} I$ for n = 23. For more explanation see Appendix B2

(B4) Angular modulation: Calculating energy distribution in side bands using Bessel functions

Eqn. (7) describing angular modulator, can be rewritten in:

$$y_t = A\sin\left(\omega\ t + \frac{m}{p}\sin pt\right) = A\left[\sin\omega\ t\cos(\frac{m}{p}\sin pt) + \cos\omega\ t\sin(\frac{m}{p}\sin pt)\right]$$

Eqn. B15

Both the terms $\cos(\frac{m}{p}\sin pt)$ and $\sin(\frac{m}{p}\sin pt)$ can be expanded in a Bessel function series:

$$\cos(\frac{m}{p}\sin pt) = J_0\left(\frac{m}{p}\right) + 2\left[J_2\left(\frac{m}{p}\right)\cos 2pt + J_4\left(\frac{m}{p}\right)\cos 4pt + \cdots\right] = \sum_{n=-\infty}^{\infty} J_{2n}\left(\frac{m}{p}\right)\cos 2npt$$

Eqn. B16

and

$$\sin(\frac{m}{p}\sin pt) = 2\left[J_1\left(\frac{m}{p}\right)\sin pt + J_3\left(\frac{m}{p}\right)\sin 3pt + \cdots\right] = \sum_{n=-\infty}^{\infty} J_{2n-1}\left(\frac{m}{p}\right)\sin(2n-1)pt$$

Eqn. B17

These functions are Bessel functions of the first kind, where the index specifies the order. Substitution of Eqn. B16 and Eqn. B17 into Eqn. B15, (with $2\sin P\cos Q = \sin(P+Q) + \sin(P+Q)$), results in

$$\begin{aligned} y_t = A.[&J_0\left(\frac{m}{p}\right)\sin\omega\ t + J_1\left(\frac{m}{p}\right)\sin(\omega\ + p)t - J_1\left(\frac{m}{p}\right)\sin(\omega\ - p)t + \\ &+ J_2\left(\frac{m}{p}\right)\sin(\omega\ + 2p)t + J_2\left(\frac{m}{p}\right)\sin(\omega\ - 2p)t + \\ &+ J_3\left(\frac{m}{p}\right)\sin(\omega\ + 3p)t - J_3\left(\frac{m}{p}\right)\sin(\omega\ - 3p)t\cdots] = A\sum_{n=-\infty}^{\infty} J_n\left(\frac{m}{p}\right)\sin(\omega\ + np)t \end{aligned}$$

Eqn. B18

The summation follows from the formula $J_{-n}(x) = (-1)^n J_n(x)$.

Parseval's theorem can be restated for angular modulations:

$$\left[J_0\left(\frac{m}{p}\right)\right]^2 + 2\left[J_1\left(\frac{m}{p}\right)\right]^2 + 2\left[J_2\left(\frac{m}{p}\right)\right]^2 + \ldots = \sum_{n=-\infty}^{\infty} J_n^{\ 2}\left(\frac{m}{p}\right) = 1$$

Eqn. B19

This rule implies that the power of the angular modulated carrier is independent on the actual modulation, which corresponds to the constant amplitude of the angular modulated carrier. The modulation redistributes energy from the carrier frequency to the many sidebands.

(B5) Progressive modulation and combinations of modulation

We used the following definition of progressive FM:

$$y(t) = A\cdot\cos\left(\left[1 + n\cdot y_{FM}\right]\cdot 2\cdot\pi\cdot\left(\frac{t}{T_C}\right)^N\right) \qquad N > 1$$

Eqn. B20

The expressions for AM (Eqn. 4), FM (Eqn. 7) and progressive FM (Eqn. B20) can be combined in many ways. The time signal of the example used in Fig. 3e is given by the expression:

$$\begin{aligned} y(t) = {} & A_1 \cdot \left(1 + m_1 \cdot y_{AM1}\right) \cdot \cos\left(\omega_{C1} \cdot t\right) + && (AM) \\ & A_2 \cdot \cos\left(\left[1 + n_1 \cdot y_{FM1}\right] \cdot \omega_{C2} \cdot t\right) + && (FM) \\ & A_3 \cdot \cos\left(\left[1 + n_2 \cdot y_{FM2}\right] \cdot 2 \cdot \pi \cdot \left(\frac{t}{T_{C3}}\right)^N\right) + && (progressive\ FM) \\ & A_4 \cdot \left(1 + m_2 \cdot y_{AM2}\right) \cdot \cos\left(\left[1 + n_3 \cdot y_{FM3}\right] \cdot \omega_{C4} \cdot t\right) && (AM\ and\ FM\ combined) \end{aligned} \quad \text{Eqn. B21}$$

The signal is composed of: (1) an amplitude-modulated signal, resulting in a carrier- and two side frequencies, (2) a frequency-modulated signal with a constant modulation-index, (3) a frequency-modulated signal with increasing modulation-amplitude (progressive FM), and (4) a combination of AM and progressive FM.

(B6) AM to PM conversions in nonlinear oscillators

We will illustrate this AM to PM conversion for an electric circuit, i.e. a nonlinear-RLC series circuit that is voltage excited. A second order differential equation describes the system dynamics:

$$L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{1}{C_{(Q)}}Q = U\cos\omega\, t\ , \qquad \text{Eqn. B22}$$

where L is the inductance (Henry) in the circuit, Q is the charge (Coulomb), R is the resistance (Ohm) and $C_{(Q)}$ is the capacitance (Farad). The input function is a sinusoidal time function ($U\cos\omega t$) in Volts. We consider the capacitance (C) in the circuit to be a function of its charge (Q), which is a common kind of non-linearity in RLC circuits. We rewrite Eqn. B22 to:

$$\frac{d^2Q}{dt^2} + \frac{R}{L}\frac{dQ}{dt} + \omega_0{}^2 F_{(Q)} = \frac{U}{L}\cos\omega\, t = B\cos\omega\, t \qquad F_{(Q)} = Q + \lambda\, Q^3 \qquad \frac{R}{L} = \omega_0 d \quad \text{Eqn. B23}$$

Differential equations with this 3rd order nonlinearity are known as Duffing equations (Duffing, 1918).

We consider the steady state solution of Eqns. B23 of the form ($Q = A\cos(\omega t + \varphi)$. The first and second derivatives of Q are: $\frac{dQ}{dt} = -\omega\, A\sin\left(\omega\, t + \varphi\right)$ and $\frac{d^2Q}{dt^2} = -\omega^2 A\cos\left(\omega\, t + \varphi\right)$.

$$Q^3 = A^3\cos^3\left(\omega\, t + \varphi\right) = \frac{3}{4}A^3\cos\left(\omega\, t + \varphi\right) + \frac{1}{4}A^3\cos 3\left(\omega\, t + \varphi\right) \qquad \text{Eqn. B24}$$

Substitution of the expression for Q in Eqn. B24 gives:

$$\left[\left(\omega_0^2 - \omega^2\right)\cos\left(\omega\, t + \varphi\right) - \omega\,\omega_0 dA\sin\left(\omega\, t + \varphi\right)\right]A + \left[3\omega_0^2\cos\left(\omega\, t + \varphi\right) + \omega_0^2\cos 3\left(\omega\, t + \varphi\right)\right]\frac{\lambda A^3}{4} = B\cos\omega\, t$$

Due to the 3^{d} order nonlinearity a third harmonic is generated. We will concentrate only on the terms concerning the fundamental ω_0 to show the AM to PM conversion. Grouping equal powers of A gives:

$$\left[(\omega_0^2 - \omega^2)^2 + d^2\omega_0^2\omega^2\right]A^2 + \frac{9}{16}\omega_0^4\lambda\ {}^2A^6 + \frac{3}{2}(\omega_0^2 - \omega^2)\omega_0^2\lambda\ A^4 = B^2 \qquad \text{Eqn. B25}$$

This expression gives the amplitude as a function of the excitation. The solution of this polynomial of the 3rd degree in A^2 can be obtained numerically. The expression can be rewritten as $\omega = f_{A,B}$:

$$\omega_0^4 A^2 + \omega^4 A^2 - 2\omega_0^2\omega^2 A^2 + d^2\omega_0^2\omega\ {}^2A^2 + \frac{3}{2}\omega_0^4\lambda\ A^4 - \frac{3}{2}\omega_0^2\omega\ {}^2\lambda\ A^4 + \frac{9}{16}\omega\ {}_0^4\lambda\ {}^2A^6 - B^2 = 0$$

$$\omega^4 - 2\omega_0{}^2\left(1 - \frac{d^2}{2} + \frac{3}{4}\lambda\ A^2\right)\omega^2 + \omega_0^4\left(1 + \frac{3}{4}\lambda\ A^2\right)^2 - \left(\frac{B}{A}\right)^2 = 0 \qquad \text{Eqn. B26}$$

The solution for is ω^2:

$$\omega^2{}_{12} = \omega_0^2\left[\left(1 - \frac{d^2}{2} + \frac{3}{4}\lambda\ A^2\right) \pm \sqrt{\left(\frac{B}{\omega_0^2 A}\right)^2 - \left(1 + \frac{3}{4}\lambda\ A^2\right)^2 + \left(1 - \frac{d^2}{2} + \frac{3}{4}\lambda\ A^2\right)^2}\right] \qquad \text{Eqn. B27}$$

Discarding higher powers of λ leads to:

$$\omega^2{}_{12} = \omega_0^2\left[\left(1 - \frac{d^2}{2} + \frac{3}{4}\lambda\ A^2\right) \pm \sqrt{\left(\frac{B}{\omega_0^2 A}\right)^2 - d^2\left(1 + \frac{3}{4}\lambda\ A^2\right) + \frac{d^4}{4}}\right] \qquad \text{Eqn. B28}$$

Only real solutions are relevant. Figure 4**a** shows ω^2/ω_0 as function of A, with B is varied from 1 to 10, λ = 0.01 and d = 0.1. Now the phase of the oscillation charge can be determined relative to the excitation (Fig. 4**b**).

$$(\omega_0^2 - \omega^2)bA - \omega_0\omega\ adA + \frac{3}{4}\lambda\ bA^3 = 0 = (\omega_0^2 - \omega^2)\sin\varphi\ A - \omega_0\omega\ cos\varphi\ dA + \frac{3}{4}\lambda\ \sin\varphi\ A^3$$

$$\tan\varphi = \frac{\omega_0\omega\ \ d}{(\omega_0^2 - \omega^2) + \frac{3}{4}\lambda\ \ A^2} \qquad \text{Eqn. B29}$$

The phase of the charge oscillation is strongly dependent on the excitation amplitude, especially around the frequencies in the neighbourhood of the low amplitude resonance frequency. The AM to PM conversion is evident even in the simplest of all nonlinear differential equation. Other peculiar properties of this type of differential equation, such as amplitude instability regions, are beyond the scope of this article.